BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 29, 2002, 18:10:12 ; Search time 2037.65 Seconds
(without alignments)
297.828 Million cell updates/sec

Title: US-09-310-844C-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 843946

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
c 1	14.4	49.7	100	5	AF174506	AF174506 Bufo garg
c 2	13.2	45.5	100	5	AF174511	AF174511 Bufo mela
c 3	13.2	45.5	100	5	AF174512	AF174512 Bufo mela
c 4	13.2	45.5	100	5	AF174513	AF174513 Bufo mela
c 5	13.2	45.5	100	5	AF174514	AF174514 Bufo mela
c 6	13.2	45.5	100	5	AF174515	AF174515 Bufo mela
c 7	13.2	45.5	100	5	AF174516	AF174516 Bufo mela
c 8	13.2	45.5	100	5	AF174517	AF174517 Bufo mela
c 9	13.2	45.5	100	5	AF174518	AF174518 Bufo mela
c 10	13.2	45.5	100	5	AF174519	AF174519 Bufo mela
c 11	13.2	45.5	100	5	AF174520	AF174520 Bufo mela
c 12	13.2	45.5	100	5	AF174521	AF174521 Bufo mela
c 13	13	44.8	82	3	AF144669	AF144669 Patella v
c 14	12.8	44.1	53	10	MDTRVNJB	X63580 M.domesticu
c 15	12.8	44.1	100	5	AF174501	AF174501 Bufo hima
c 16	12.8	44.1	100	5	AF174505	AF174505 Bufo garg
c 17	12.8	44.1	100	5	AF174507	AF174507 Bufo garg
c 18	12.8	44.1	100	5	AF174508	AF174508 Bufo bank
19	12.2	42.1	44	6	AX008706	AX008706 Sequence
c 20	12.2	42.1	45	6	AX008707	AX008707 Sequence
21	12.2	42.1	69	6	AR052906	AR052906 Sequence
22	12.2	42.1	69	6	AR054269	AR054269 Sequence
23	12.2	42.1	69	6	AR054471	AR054471 Sequence
c 24	12.2	42.1	70	6	A42881	A42881 Sequence 13
c 25	12.2	42.1	73	6	E02131	E02131 Pseudoknot
c 26	12	41.4	24	6	AX291705	AX291705 Sequence
27	11.8	40.7	21	6	AR080204	AR080204 Sequence
28	11.8	40.7	21	6	AX088729	AX088729 Sequence
c 29	11.8	40.7	21	6	AX088730	AX088730 Sequence
c 30	11.8	40.7	33	10	MDTRVNJK	X63589 M.domesticu
c 31	11.8	40.7	40	10	MDTRVNJA	X63579 M.domesticu
c 32	11.8	40.7	46	10	MDTRVNJC	X63581 M.domesticu
33	11.8	40.7	47	6	A82690	A82690 Sequence 35
34	11.8	40.7	47	6	A82705	A82705 Sequence 50
c 35	11.8	40.7	51	10	MDTRVNJD	X63582 M.domesticu
c 36	11.8	40.7	68	7	PT7CLS	M11570 Bacteriopha
c 37	11.8	40.7	76	10	MUSG1J	M26768 Mus musculu
c 38	11.8	40.7	77	6	AX092950	AX092950 Sequence
39	11.8	40.7	77	6	AX092951	AX092951 Sequence
c 40	11.8	40.7	77	6	AX167320	AX167320 Sequence
41	11.8	40.7	77	6	AX167321	AX167321 Sequence
c 42	11.8	40.7	82	3	SASHOXW20H	L14920 Saccoglossu
c 43	11.8	40.7	100	5	AF174502	AF174502 Bufo andr
c 44	11.6	40.0	36	6	AR176463	AR176463 Sequence
45	11.6	40.0	36	6	AR176468	AR176468 Sequence

ALIGNMENTS

RESULT 1
AF174506/c
LOCUS AF174506 100 bp DNA linear VRT 20-APR-2000
DEFINITION Bufo gargarizans KIZ-97F001 cytochrome b gene, partial cds; mitochondrial gene for mitochondrial product.
ACCESSION AF174506
VERSION AF174506.1 GI:7620449
KEYWORDS .
SOURCE Bufo gargarizans.
ORGANISM Mitochondrion Bufo gargarizans
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae; Bufo.
REFERENCE 1 (bases 1 to 100)
AUTHORS Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
TITLE Phylogeny of East Asian bufonids inferred from mitochondrial DNA sequences (Anura: Amphibia)
JOURNAL Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
MEDLINE 20179527
PUBMED 10712847
REFERENCE 2 (bases 1 to 100)
AUTHORS Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
TITLE Direct Submission

```
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo gargarizans"
                     /organelle="mitochondrion"
                     /specimen_voucher="KIZ-97F001"
                     /db_xref="taxon:30331"
                     /country="China: Fujian, Chun-an Co."
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64625.1"
                     /db_xref="GI:7620450"
                     /translation="MIGQLASISYFSIFIIFIPLLGLTENKLTQLS"
BASE COUNT       27 a     33 c     11 g     29 t
ORIGIN

  Query Match             49.7%;  Score 14.4;  DB 5;  Length 100;
  Best Local Similarity   45.5%;  Pred. No. 5.1e+02;
  Matches   10;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      6 uncuuunnguaagcccnangng 27
          :  :::  |:|||||| | | |
Db     80 TGTTTTCTGTAAGCCCGAGGAG 59


RESULT   2
AF174511/c
LOCUS       AF174511                100 bp    DNA     linear   VRT 20-APR-2000
DEFINITION  Bufo melanostictus KIZ-95L001 cytochrome b gene, partial cds;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF174511
VERSION     AF174511.1  GI:7620459
KEYWORDS    .
SOURCE      Bufo melanostictus.
  ORGANISM  Mitochondrion Bufo melanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;
            Bufo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
  TITLE     Phylogeny of East Asian bufonids inferred from mitochondrial DNA
            sequences (Anura: Amphibia)
  JOURNAL   Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
  MEDLINE   20179527
   PUBMED   10712847
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo melanostictus"
                     /organelle="mitochondrion"
                     /specimen_voucher="KIZ-95L001"
                     /db_xref="taxon:30335"
                     /country="China: Yunnan, Pingbian Co."
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64630.1"
                     /db_xref="GI:7620460"
                     /translation="MIGQLASISYFTIFIIFIPLLGLTENNLARLN"
BASE COUNT       28 a     29 c     10 g     33 t
ORIGIN


  Query Match             45.5%;  Score 13.2;  DB 5;  Length 100;
  Best Local Similarity   41.7%;  Pred. No. 3.1e+03;
  Matches   10;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          | :  :::  |:||| || | | |
Db     82 GTTATTTTCTGTAAGTCCTAAGAG 59


RESULT   3
AF174512/c
LOCUS       AF174512                100 bp    DNA     linear   VRT 20-APR-2000
DEFINITION  Bufo melanostictus KIZ-97H001 cytochrome b gene, partial cds;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF174512
VERSION     AF174512.1  GI:7620461
KEYWORDS    .
SOURCE      Bufo melanostictus.
  ORGANISM  Mitochondrion Bufo melanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;
            Bufo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
  TITLE     Phylogeny of East Asian bufonids inferred from mitochondrial DNA
            sequences (Anura: Amphibia)
  JOURNAL   Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
  MEDLINE   20179527
   PUBMED   10712847
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo melanostictus"
                     /organelle="mitochondrion"
                     /specimen_voucher="KIZ-97H001"
                     /db_xref="taxon:30335"
                     /country="China: Hainan, Lingshui Co."
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64631.1"
                     /db_xref="GI:7620462"
                     /translation="MIGQLASISYFAIFIVFIPLLGLTENNLARLN"
BASE COUNT       25 a     31 c     13 g     31 t
ORIGIN


  Query Match             45.5%;  Score 13.2;  DB 5;  Length 100;
  Best Local Similarity   41.7%;  Pred. No. 3.1e+03;
  Matches   10;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          | :  :::  |: |||| | | |
Db     82 GTTGTTTTCTGTGAGCCCTAAGAG 59


RESULT   4
AF174513/c
LOCUS       AF174513                100 bp    DNA     linear   VRT 20-APR-2000
DEFINITION  Bufo melanostictus KIZ-97L202 cytochrome b gene, partial cds;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF174513
VERSION     AF174513.1  GI:7620463
```

```
KEYWORDS    .
SOURCE      Bufo melanostictus.
  ORGANISM  Mitochondrion Bufo melanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;
            Bufo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
  TITLE     Phylogeny of East Asian bufonids inferred from mitochondrial DNA
            sequences (Anura: Amphibia)
  JOURNAL   Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
  MEDLINE   20179527
   PUBMED   10712847
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo melanostictus"
                     /organelle="mitochondrion"
                     /specimen_voucher="KIZ-97L202"
                     /db_xref="taxon:30335"
                     /country="China: Yunnan, Lunan Co."
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64632.1"
                     /db_xref="GI:7620464"
                     /translation="MIGQLASISYFTIFIIFIPLLGLTENNLARLN"
BASE COUNT       28 a     29 c     10 g     33 t
ORIGIN


  Query Match             45.5%;  Score 13.2;  DB 5;  Length 100;
  Best Local Similarity   41.7%;  Pred. No. 3.1e+03;
  Matches   10;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          | :  :::  |:||| || | | |
Db     82 GTTATTTTCTGTAAGTCCTAAGAG 59


RESULT   5
AF174514/c
LOCUS       AF174514                100 bp    DNA     linear   VRT 20-APR-2000
DEFINITION  Bufo melanostictus KIZ-97L118 cytochrome b gene, partial cds;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF174514
VERSION     AF174514.1  GI:7620465
KEYWORDS    .
SOURCE      Bufo melanostictus.
  ORGANISM  Mitochondrion Bufo melanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;
            Bufo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
  TITLE     Phylogeny of East Asian bufonids inferred from mitochondrial DNA
            sequences (Anura: Amphibia)
  JOURNAL   Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
  MEDLINE   20179527
   PUBMED   10712847
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo melanostictus"
                     /organelle="mitochondrion"
                     /specimen_voucher="KIZ-97L118"
                     /db_xref="taxon:30335"
                     /country="China: Yunnan, Mengla Co."
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64633.1"
                     /db_xref="GI:7620466"
                     /translation="MIGQLASISYFTIFIIFIPLLGLTENNLARLN"
BASE COUNT       28 a     29 c     10 g     33 t
ORIGIN


  Query Match             45.5%;  Score 13.2;  DB 5;  Length 100;
  Best Local Similarity   41.7%;  Pred. No. 3.1e+03;
  Matches   10;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          | :  :::  |:||| || | | |
Db     82 GTTATTTTCTGTAAGTCCTAAGAG 59


RESULT   6
AF174515/c
LOCUS       AF174515                100 bp    DNA     linear   VRT 20-APR-2000
DEFINITION  Bufo melanostictus KIZ-92L080 cytochrome b gene, partial cds;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF174515
VERSION     AF174515.1  GI:7620467
KEYWORDS    .
SOURCE      Bufo melanostictus.
  ORGANISM  Mitochondrion Bufo melanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;
            Bufo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
  TITLE     Phylogeny of East Asian bufonids inferred from mitochondrial DNA
            sequences (Anura: Amphibia)
  JOURNAL   Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
  MEDLINE   20179527
   PUBMED   10712847
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo melanostictus"
                     /organelle="mitochondrion"
                     /specimen_voucher="KIZ-92L080"
                     /db_xref="taxon:30335"
                     /country="China: Yunnan, Tengchong Co."
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64634.1"
                     /db_xref="GI:7620468"
                     /translation="MIGQLASISYFTIFIIFIPLLGLTENNLARLN"
BASE COUNT       28 a     29 c     10 g     33 t
ORIGIN
```

```
  Query Match            45.5%;  Score 13.2;  DB 5;  Length 100;
  Best Local Similarity  41.7%;  Pred. No. 3.1e+03;
  Matches   10;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          | :  :::  |:||| || | | |
Db     82 GTTATTTTCTGTAAGTCCTAAGAG 59


RESULT   7
AF174516/c
LOCUS       AF174516                100 bp    DNA     linear   VRT 20-APR-2000
DEFINITION  Bufo melanostictus KIZ-97L372 cytochrome b gene, partial cds;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF174516
VERSION     AF174516.1  GI:7620469
KEYWORDS    .
SOURCE      Bufo melanostictus.
  ORGANISM  Mitochondrion Bufo melanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;
            Bufo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
  TITLE     Phylogeny of East Asian bufonids inferred from mitochondrial DNA
            sequences (Anura: Amphibia)
  JOURNAL   Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
  MEDLINE   20179527
   PUBMED   10712847
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo melanostictus"
                     /organelle="mitochondrion"
                     /specimen_voucher="KIZ-97L372"
                     /db_xref="taxon:30335"
                     /country="Taiwan"
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64635.1"
                     /db_xref="GI:7620470"
                     /translation="MIGQLASISYFAIFIVFIPLLGLTENNLARLN"
BASE COUNT       25 a     30 c     13 g     32 t
ORIGIN


  Query Match            45.5%;  Score 13.2;  DB 5;  Length 100;
  Best Local Similarity  41.7%;  Pred. No. 3.1e+03;
  Matches   10;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          | :  :::  |: ||||| | | |
Db     82 GTTGTTTTCTGTGAGCCCTAAGAG 59


RESULT   8
AF174517/c
LOCUS       AF174517                100 bp    DNA     linear   VRT 20-APR-2000
DEFINITION  Bufo melanostictus ROM 33162 cytochrome b gene, partial cds;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF174517
VERSION     AF174517.1  GI:7620471
KEYWORDS    .
SOURCE      Bufo melanostictus.
  ORGANISM  Mitochondrion Bufo melanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;
            Bufo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
  TITLE     Phylogeny of East Asian bufonids inferred from mitochondrial DNA
            sequences (Anura: Amphibia)
  JOURNAL   Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
  MEDLINE   20179527
   PUBMED   10712847
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo melanostictus"
                     /organelle="mitochondrion"
                     /specimen_voucher="ROM 33162"
                     /db_xref="taxon:30335"
                     /country="Viet Nam: Gia Lia, Ko Rong"
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64636.1"
                     /db_xref="GI:7620472"
                     /translation="MIGQLASISYFAIFIVFIPLLGLTENNLARLN"
BASE COUNT       25 a     30 c     13 g     32 t
ORIGIN


  Query Match            45.5%;  Score 13.2;  DB 5;  Length 100;
  Best Local Similarity  41.7%;  Pred. No. 3.1e+03;
  Matches   10;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          | :  :::  |: ||||| | | |
Db     82 GTTGTTTTCTGTGAGCCCTAAGAG 59


RESULT   9
AF174518/c
LOCUS       AF174518                100 bp    DNA     linear   VRT 20-APR-2000
DEFINITION  Bufo melanostictus ROM 32540 cytochrome b gene, partial cds;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF174518
VERSION     AF174518.1  GI:7620473
KEYWORDS    .
SOURCE      Bufo melanostictus.
  ORGANISM  Mitochondrion Bufo melanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;
            Bufo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
  TITLE     Phylogeny of East Asian bufonids inferred from mitochondrial DNA
            sequences (Anura: Amphibia)
  JOURNAL   Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
  MEDLINE   20179527
   PUBMED   10712847
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo melanostictus"
                     /organelle="mitochondrion"
```

```
                     /specimen_voucher="ROM 32540"
                     /db_xref="taxon:30335"
                     /country="Viet Nam: Gia Lia, Ko Rong"
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64637.1"
                     /db_xref="GI:7620474"
                     /translation="MIGQLASISYFAIFIVFIPLLGLTENNLARLN"
BASE COUNT       25 a     30 c     13 g     32 t
ORIGIN


  Query Match             45.5%;  Score 13.2;  DB 5;  Length 100;
  Best Local Similarity   41.7%;  Pred. No. 3.1e+03;
  Matches   10;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnangng 27
           | :  :::  |: ||||| | | |
Db      82 GTTGTTTTCTGTGAGCCCTAAGAG 59


RESULT  10
AF174519/c
LOCUS       AF174519                 100 bp    DNA     linear   VRT 20-APR-2000
DEFINITION  Bufo melanostictus ROM 33163 cytochrome b gene, partial cds;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF174519
VERSION     AF174519.1  GI:7620475
KEYWORDS    .
SOURCE      Bufo melanostictus.
  ORGANISM  Mitochondrion Bufo melanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;
            Bufo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
  TITLE     Phylogeny of East Asian bufonids inferred from mitochondrial DNA
            sequences (Anura: Amphibia)
  JOURNAL   Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
  MEDLINE   20179527
   PUBMED   10712847
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo melanostictus"
                     /organelle="mitochondrion"
                     /specimen_voucher="ROM 33163"
                     /db_xref="taxon:30335"
                     /country="Viet Nam: Gia Lia, Ko Rong"
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64638.1"
                     /db_xref="GI:7620476"
                     /translation="MIGQLASISYFAIFIVFIPLLGLTENNLARLN"
BASE COUNT       25 a     30 c     13 g     32 t
ORIGIN


  Query Match             45.5%;  Score 13.2;  DB 5;  Length 100;
  Best Local Similarity   41.7%;  Pred. No. 3.1e+03;
  Matches   10;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnangng 27
           | :  :::  |: ||||| | | |
Db      82 GTTGTTTTCTGTGAGCCCTAAGAG 59


RESULT  11
AF174520/c
LOCUS       AF174520                 100 bp    DNA     linear   VRT 20-APR-2000
DEFINITION  Bufo melanostictus ROM 33855 cytochrome b gene, partial cds;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF174520
VERSION     AF174520.1  GI:7620477
KEYWORDS    .
SOURCE      Bufo melanostictus.
  ORGANISM  Mitochondrion Bufo melanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;
            Bufo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
  TITLE     Phylogeny of East Asian bufonids inferred from mitochondrial DNA
            sequences (Anura: Amphibia)
  JOURNAL   Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
  MEDLINE   20179527
   PUBMED   10712847
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo melanostictus"
                     /organelle="mitochondrion"
                     /specimen_voucher="ROM 33855"
                     /db_xref="taxon:30335"
                     /country="Viet Nam: Dac Lac, Yok Don"
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64639.1"
                     /db_xref="GI:7620478"
                     /translation="MIGQLASISYFAIFIVFIPLLGLTENNLARLN"
BASE COUNT       25 a     30 c     13 g     32 t
ORIGIN


  Query Match             45.5%;  Score 13.2;  DB 5;  Length 100;
  Best Local Similarity   41.7%;  Pred. No. 3.1e+03;
  Matches   10;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnangng 27
           | :  :::  |: ||||| | | |
Db      82 GTTGTTTTCTGTGAGCCCTAAGAG 59


RESULT  12
AF174521/c
LOCUS       AF174521                 100 bp    DNA     linear   VRT 20-APR-2000
DEFINITION  Bufo melanostictus ROM 33861 cytochrome b gene, partial cds;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF174521
VERSION     AF174521.1  GI:7620479
KEYWORDS    .
SOURCE      Bufo melanostictus.
  ORGANISM  Mitochondrion Bufo melanostictus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;
            Bufo.
```

```
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
  TITLE     Phylogeny of East Asian bufonids inferred from mitochondrial DNA
            sequences (Anura: Amphibia)
  JOURNAL   Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
  MEDLINE   20179527
   PUBMED   10712847
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo melanostictus"
                     /organelle="mitochondrion"
                     /specimen_voucher="ROM 33861"
                     /db_xref="taxon:30335"
                     /country="Viet Nam: Dac Lac, Yok Don"
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64640.1"
                     /db_xref="GI:7620480"
                     /translation="MIGQLASISYFAIFIVFIPLLGLTENNLARLN"
BASE COUNT       25 a     30 c     13 g     32 t
ORIGIN
```

Query Match 45.5%; Score 13.2; DB 5; Length 100;
Best Local Similarity 41.7%; Pred. No. 3.1e+03;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

```
Qy       4 gauncuuunnguaagcccnangng 27
           | :  :::  |: ||||| | | |
Db      82 GTTGTTTTCTGTGAGCCCTAAGAG 59
```

```
RESULT  13
AF144669/c
LOCUS       AF144669                 82 bp    DNA     linear   INV 04-AUG-1999
DEFINITION  Patella vulgata antennapedia-like homeodomain protein HB3 gene,
            partial cds.
ACCESSION   AF144669
VERSION     AF144669.1  GI:5690267
KEYWORDS    .
SOURCE      common limpet.
  ORGANISM  Patella vulgata
            Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
            Patelloidea; Patellidae; Patella.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   de Rosa,R., Grenier,J.K., Andreeva,T., Cook,C.E., Adoutte,A.,
            Akam,M., Carroll,S.B. and Balavoine,G.
  TITLE     Hox genes in brachiopods and priapulids and protostome evolution
  JOURNAL   Nature 399 (6738), 772-776 (1999)
  MEDLINE   99318125
   PUBMED   10391241
REFERENCE   2  (bases 1 to 82)
  AUTHORS   de Rosa,R., Lartillot,N. and Adoutte,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-APR-1999) Centre de Genetique Moleculaire, Avenue de
            la Terrasse, Gif-sur-Yvette 91198, France
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Patella vulgata"
                     /db_xref="taxon:6465"
     mRNA            <1. .>82
                     /product="antennapedia-like homeodomain protein HB3"
     CDS             <1. .>82
                     /codon_start=2
                     /product="antennapedia-like homeodomain protein HB3"
                     /protein_id="AAD47008.1"
                     /db_xref="GI:5690268"
                     /translation="HFNRYLTRRRRIEVAHMLGLTERQIKI"
BASE COUNT       34 a     16 c     17 g     15 t
ORIGIN
```

Query Match 44.8%; Score 13; DB 3; Length 82;
Best Local Similarity 56.2%; Pred. No. 4.1e+03;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```
Qy       8 cuuunnguaagcccna 23
           |:::  |:|||||| |
Db      69 CTTTCTGTAAGCCCGA 54
```

```
RESULT  14
MDTRVNJB/c
LOCUS       MDTRVNJB                 53 bp    mRNA    linear   ROD 07-MAR-1993
DEFINITION  M.domesticus DBA/2 rearranged T-cell receptor (Vgamma2-N-Jgamma2).
ACCESSION   X63580
VERSION     X63580.1  GI:57892
KEYWORDS    joining region; N-region; T-cell receptor; variable region.
SOURCE      western European house mouse.
  ORGANISM  Mus musculus domesticus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   Roger,T.R.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 53)
  AUTHORS   Roger,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-1991) T. Roger, Laboratoire
            d'Immunodifferenciation, Service du Pr SEMAN, Institut J.MONOD, 2,
            Place JUSSIEU, 75251 PARIS CEDEX 05, FRANCE
FEATURES             Location/Qualifiers
     source          1. .53
                     /organism="Mus musculus domesticus"
                     /strain="DBA/2"
                     /db_xref="taxon:10092"
                     /chromosome="13"
                     /dev_stage="Seed, cell expansion stage"
                     /tissue_type="Spleen"
                     /cell_type="CD4- CD8- T Lymphocyte clone"
                     /cell_line="1G3"
                     /clone_lib="library lambda gt11"
                     /clone="1G3gamma5"
     misc_feature    1. .24
                     /product="End of the T-cell receptor Vgamma2 gene
                     segment"
     misc_feature    25. .31
                     /product="N Sequence"
     misc_feature    32. .53
                     /product="Begining of the T cell receptor Jgamma2  gene
                     segment"
BASE COUNT       10 a     15 c     11 g     17 t
ORIGIN
```

Query Match 44.1%; Score 12.8; DB 10; Length 53;
Best Local Similarity 52.4%; Pred. No. 5.3e+03;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```
Qy       5 auncuuunnguaagcccnang 25
           |: | ::  | ||||||| | |
Db      51 ATACCTTGTGAAAGCCCGAGG 31
```

```
RESULT  15
AF174501/c
LOCUS       AF174501                100 bp    DNA     linear   VRT 20-APR-2000
```

```
DEFINITION  Bufo himalayanus KIZ-95L010 cytochrome b gene, partial cds;
            mitochondrial gene for mitochondrial product.
ACCESSION   AF174501
VERSION     AF174501.1  GI:7620439
KEYWORDS    .
SOURCE      Bufo himalayanus.
  ORGANISM  Mitochondrion Bufo himalayanus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;
            Bufo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J., Yang,D. and Murphy,R.W.
  TITLE     Phylogeny of East Asian bufonids inferred from mitochondrial DNA
            sequences (Anura: Amphibia)
  JOURNAL   Mol. Phylogenet. Evol. 14 (3), 423-435 (2000)
  MEDLINE   20179527
   PUBMED   10712847
REFERENCE   2  (bases 1 to 100)
  AUTHORS   Liu,W., Lathrop,A., Fu,J. and Murphy,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JUL-1999) CBCB, Royal Ontario Museum, 100 Queen's
            Park, Toronto, Ontario M5S 2C6, Canada
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Bufo himalayanus"
                     /organelle="mitochondrion"
                     /specimen_voucher="KIZ-95L010"
                     /db_xref="taxon:103434"
                     /country="China: Yunnan, Gongshan Co."
     mRNA            <1. .>100
                     /product="cytochrome b"
     CDS             <1. .100
                     /codon_start=2
                     /transl_table=2
                     /product="cytochrome b"
                     /protein_id="AAF64620.1"
                     /db_xref="GI:7620440"
                     /translation="MIGQLASISYFTIFIIFIPLLGLTENKLARLN"
BASE COUNT       28 a     31 c     11 g     30 t
ORIGIN


  Query Match           44.1%;  Score 12.8;  DB 5;  Length 100;
  Best Local Similarity 40.9%;  Pred. No. 5.6e+03;
  Matches    9;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       6 uncuuunnguaagcccnangng 27
           :  :::  |: ||||| | | |
        80 TGTTTTCTGTCAGCCCTAGGAG 59


Search completed: April 29, 2002, 22:07:08
Job time: 14216 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 29, 2002, 21:37:54 ; Search time 310.59 Seconds
(without alignments)
160.309 Million cell updates/sec

Title: US-09-310-844C-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 2046006

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	18	62.1	29	21	AAA70827	Molecular interact
2	18	62.1	29	21	AAA70828	Molecular interact
3	18	62.1	29	21	AAA70829	Molecular interact
4	18	62.1	29	21	AAA70830	Molecular interact
5	18	62.1	42	21	AAA71113	Molecular interact
6	18	62.1	42	21	AAA71114	Molecular interact
7	18	62.1	42	21	AAA71115	Molecular interact
8	18	62.1	42	21	AAA71116	Molecular interact
9	18	62.1	42	21	AAA71118	Molecular interact
10	18	62.1	42	21	AAA71119	Molecular interact
11	18	62.1	42	21	AAA71120	Molecular interact
12	18	62.1	42	21	AAA71121	Molecular interact
13	18	62.1	42	21	AAA71123	Molecular interact
14	18	62.1	42	21	AAA71124	Molecular interact
15	18	62.1	42	21	AAA71126	Molecular interact
16	18	62.1	42	21	AAA71127	Molecular interact
17	18	62.1	42	21	AAA71128	Molecular interact
18	18	62.1	42	21	AAA71129	Molecular interact
19	18	62.1	42	21	AAA71131	Molecular interact
20	18	62.1	42	21	AAA71132	Molecular interact
21	18	62.1	44	21	AAA71112	Molecular interact
22	18	62.1	44	21	AAA71125	Molecular interact
23	18	62.1	44	21	AAA71133	Molecular interact
24	18	62.1	45	21	AAA70824	Molecular interact
25	18	62.1	45	21	AAA70825	Molecular interact
26	18	62.1	45	21	AAA70826	Molecular interact
27	18	62.1	46	21	AAA71085	Molecular interact
28	18	62.1	46	21	AAA71087	Molecular interact
29	18	62.1	46	21	AAA71088	Molecular interact
30	18	62.1	46	21	AAA71089	Molecular interact
31	18	62.1	46	21	AAA71090	Molecular interact
32	18	62.1	46	21	AAA71093	Molecular interact
33	18	62.1	46	21	AAA71094	Molecular interact
34	18	62.1	46	21	AAA71095	Molecular interact
35	18	62.1	46	21	AAA71096	Molecular interact
36	18	62.1	46	21	AAA71099	Molecular interact
37	18	62.1	46	21	AAA71100	Molecular interact
38	18	62.1	46	21	AAA71103	Molecular interact
39	18	62.1	46	21	AAA71104	Molecular interact
40	18	62.1	46	21	AAA71105	Molecular interact
41	18	62.1	46	21	AAA71106	Molecular interact
42	18	62.1	46	21	AAA71107	Molecular interact
43	18	62.1	46	21	AAA71109	Molecular interact
44	18	62.1	46	21	AAA71110	Molecular interact
45	18	62.1	46	21	AAA71111	Molecular interact

ALIGNMENTS

RESULT 1
AAA70827
ID AAA70827 standard; RNA; 29 BP.
XX
AC AAA70827;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #27.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Synthetic.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 4 A; 4 C; 5 G; 5 U; 11 other;

Query Match 62.1%; Score 18; DB 21; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.97;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||||||||||||||||||||||||
Db 4 gauncuuunnguaagcccnangng 27

RESULT 2
AAA70828
ID AAA70828 standard; RNA; 29 BP.
XX
AC AAA70828;

27-APR-2001 (first entry)

DE Molecular interaction site RNA #28.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Homo sapiens.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 5 A; 5 C; 7 G; 12 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 29;
Best Local Similarity 75.0%; Pred. No. 0.97;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 4 gauucuuuuuguaagcccuagggg 27

RESULT 3
AAA70829
ID AAA70829 standard; RNA; 29 BP.
XX
AC AAA70829;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #29.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Mus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 9 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 29;
Best Local Similarity 75.0%; Pred. No. 0.97;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 4 gauucuuuuuguaagccccaaggg 27

RESULT 4
AAA70830
ID AAA70830 standard; RNA; 29 BP.
XX
AC AAA70830;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #30.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Rattus sp.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Claim 235; Page 235; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 29 BP; 8 A; 6 C; 6 G; 9 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 29;
Best Local Similarity 75.0%; Pred. No. 0.97;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 4 gauucuuuuuguaagccccaaggg 27

RESULT 5
AAA71113
ID AAA71113 standard; RNA; 42 BP.
XX
AC AAA71113;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #189.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Example 7; Figure 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 17 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 75.0%; Pred. No. 1;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 7 gauucuuuuuguaagcccuacggg 30

RESULT 6
AAA71114
ID AAA71114 standard; RNA; 42 BP.
XX
AC AAA71114;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #190.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Example 7; Figure 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 11 A; 8 C; 7 G; 16 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 75.0%; Pred. No. 1;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 7 gauucuuuuuguaagcccuaggcg 30

RESULT 7
AAA71115
ID AAA71115 standard; RNA; 42 BP.
XX
AC AAA71115;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #191.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Example 7; Figure 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 75.0%; Pred. No. 1;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 7 gauucuuuuuguaagccccaaggg 30

RESULT 8
AAA71116
ID AAA71116 standard; RNA; 42 BP.
XX
AC AAA71116;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #192.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Example 7; Figure 122; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 75.0%; Pred. No. 1;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||| |||| |||||||| | | |
Db 7 gauucuuuuuguaagccccaaggg 30

RESULT 9
AAA71118
ID AAA71118 standard; DNA; 42 BP.
XX
AC AAA71118;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #124.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 17 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 1;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||: |::: |:|||||| | | |
Db 7 gattctttttgtaagccctacggg 30

RESULT 10
AAA71119
ID AAA71119 standard; DNA; 42 BP.
XX
AC AAA71119;

27-APR-2001 (first entry)

DE Molecular interaction site DNA #125.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 11 A; 8 C; 7 G; 16 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 1;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||: |::: |:|||||| | | |
Db 7 gattctttttgtaagccctaggcg 30

RESULT 11
AAA71120
ID AAA71120 standard; DNA; 42 BP.
XX
AC AAA71120;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #126.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 1;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||: |::: |:|||||| | | |
Db 7 gattctttttgtaagccccaaggg 30

RESULT 12
AAA71121
ID AAA71121 standard; DNA; 42 BP.
XX
AC AAA71121;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #127.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 13 A; 7 C; 7 G; 15 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 1;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||: |::: |:|||||| | | |
Db 7 gattctttttgtaagccccaaggg 30

RESULT 13
AAA71123
ID AAA71123 standard; DNA; 42 BP.
XX
AC AAA71123;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #129.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 9 A; 6 C; 9 G; 18 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 1;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||: |::: |:|||||| | | |
Db 7 gattctttttgtaagccctagggg 30

RESULT 14
AAA71124
ID AAA71124 standard; DNA; 42 BP.
XX
AC AAA71124;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site DNA #130.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Example 7; Figure 125; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 11 A; 10 C; 7 G; 14 T; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 54.2%; Pred. No. 1;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagcccnangng 27
||: |::: |:|||||| | | |
Db 7 gatcctttctgtaagccctacggg 30

RESULT 15
AAA71126
ID AAA71126 standard; RNA; 42 BP.
XX
AC AAA71126;
XX
DT 27-APR-2001 (first entry)
XX
DE Molecular interaction site RNA #195.
XX
KW Modulator; identification; molecular interaction; virtual library; ss.
XX
OS Unidentified.
XX
PN WO9958947-A2.
XX
PD 18-NOV-1999.
XX
PF 12-MAY-1999; 99WO-US10361.
XX
PR 12-MAY-1998; 98US-0076404.
PR 12-MAY-1998; 98US-0085092.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ecker DJ, Griffey R, Crooke ST, Sampath R, Swayze E, Mohan V;
PI Hofstadler S, McNeil J;
XX
DR WPI; 2000-086439/07.
XX
PT Identifying compounds which modulate activity of target biomolecules,
PT used to provide compounds which can be used as pharmacological,

PT agricultural and industrial compounds -
XX
PS Example 7; Figure 126; 405pp; English.
XX
CC This invention describes a novel method for identifying compounds which
CC modulate the activity of a target biomolecule. The method uses
CC 3-dimensional representations of the biomolecule and a library of
CC compounds and comprises (a) identifying at least one molecular
CC interaction site of the target RNA; (b) generating in silico a virtual
CC library of compounds predicted or calculated to interact with the
CC molecular interaction site; and (c) comparing 3-dimensional (3-D)
CC representations of the target RNA with members of the virtual library of
CC compounds to generate a hierarchy of the compounds ranked in accordance
CC with their respective ability to form physical interactions with the
CC molecular interaction site. The method also describes (1) RNA comprising
CC a joined sequence of at least 24 nucleotides but not more than 70
CC nucleotides and having secondary structure defined by: (a) 3 nucleotides
CC forming a first side of a first double stranded (ds) region; (b) 2
CC nucleotides forming a first side of an internal loop region; (c) 4
CC nucleotides forming a first side of a second ds region; (d) 4 or 5
CC nucleotides forming an end loop region; (e) 4 nucleotides forming a
CC second side of the second ds region; (f) 4 nucleotides forming a second
CC side of the internal loop region; and (g) 3 nucleotides forming a second
CC side of the first ds region; (2) a purified and isolated RNA fragment
CC comprising the human sequence UUUACAACAUAAUCUAGUUUACAGAAAAAUC (II). The
CC methods and products can be used for identifying agents which modulate
CC the activity of biomolecules, particularly RNA. Such agents can be used
CC as pharmaceutical, agricultural or industrial compounds.
XX
SQ Sequence 42 BP; 12 A; 7 C; 6 G; 17 U; 0 other;

Query Match 62.1%; Score 18; DB 21; Length 42;
Best Local Similarity 75.0%; Pred. No. 1;
Matches 18; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy      4 gauncuuunnguaagcccnangng 27
          ||| ||||  |||||||| | | |
Db      7 gauucuuuuuguaagcccuacggg 30
```

Search completed: April 29, 2002, 22:45:09
Job time: 4035 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 29, 2002, 19:50:29 ; Search time 60.62 Seconds
(without alignments)
117.509 Million cell updates/sec

Title: US-09-310-844C-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 613726

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	12.2	42.1	27	6	5258283-10	Patent No. 5258283
	2	12.2	42.1	69	2	US-08-410-654B-30	Sequence 30, Appl
	3	12.2	42.1	69	2	US-08-474-851-30	Sequence 30, Appl
	4	12.2	42.1	69	2	US-08-481-560-30	Sequence 30, Appl
	5	11.8	40.7	21	2	US-08-747-536-10	Sequence 10, Appl
c	6	11.6	40.0	36	4	US-08-218-369-7	Sequence 7, Appli
	7	11.6	40.0	36	4	US-08-218-369-15	Sequence 15, Appl
c	8	11.6	40.0	36	5	PCT-US95-03742-7	Sequence 7, Appli
	9	11.6	40.0	36	5	PCT-US95-03742-15	Sequence 15, Appl
c	10	11.2	38.6	25	1	US-08-741-881-28	Sequence 28, Appl
c	11	11.2	38.6	25	1	US-08-739-158-28	Sequence 28, Appl
c	12	11.2	38.6	25	2	US-08-739-167-28	Sequence 28, Appl
c	13	11.2	38.6	25	3	US-08-404-796-28	Sequence 28, Appl
c	14	11.2	38.6	25	3	US-08-931-869-28	Sequence 28, Appl
c	15	11.2	38.6	25	4	US-09-350-399-28	Sequence 28, Appl
	16	11.2	38.6	33	1	US-08-741-881-29	Sequence 29, Appl
	17	11.2	38.6	33	1	US-08-739-158-29	Sequence 29, Appl
	18	11.2	38.6	33	2	US-08-739-167-29	Sequence 29, Appl
	19	11.2	38.6	33	3	US-08-404-796-29	Sequence 29, Appl
	20	11.2	38.6	33	3	US-08-931-869-29	Sequence 29, Appl
	21	11.2	38.6	33	4	US-09-350-399-29	Sequence 29, Appl
c	22	11.2	38.6	36	2	US-08-642-045B-17	Sequence 17, Appl
c	23	11.2	38.6	36	3	US-08-852-268-17	Sequence 17, Appl
	24	11.2	38.6	70	4	US-09-364-380-29	Sequence 29, Appl
c	25	11	37.9	31	1	US-08-323-531-71	Sequence 71, Appl
c	26	11	37.9	31	1	US-08-198-094-71	Sequence 71, Appl
c	27	11	37.9	31	3	US-08-480-640A-119	Sequence 119, App
c	28	11	37.9	31	3	US-08-295-802-119	Sequence 119, App
c	29	11	37.9	31	4	US-08-107-794A-71	Sequence 71, Appl
c	30	11	37.9	31	4	US-08-488-237A-119	Sequence 119, App
c	31	11	37.9	31	4	US-08-375-992A-119	Sequence 119, App
c	32	11	37.9	31	5	PCT-US93-07424-71	Sequence 71, Appl
c	33	11	37.9	31	5	PCT-US95-02087-71	Sequence 71, Appl
	34	11	37.9	70	2	US-08-488-402A-127	Sequence 127, App
	35	11	37.9	70	2	US-08-484-552A-127	Sequence 127, App
	36	11	37.9	70	5	PCT-US96-09472-127	Sequence 127, App
	37	10.8	37.2	19	1	US-08-365-109B-1	Sequence 1, Appli
	38	10.8	37.2	19	1	US-08-365-109B-3	Sequence 3, Appli
c	39	10.8	37.2	20	4	US-09-560-594-53	Sequence 53, Appl
c	40	10.8	37.2	25	4	US-08-943-731-336	Sequence 336, App
	41	10.8	37.2	27	4	US-09-253-396A-137	Sequence 137, App
c	42	10.8	37.2	45	1	US-08-171-389-130	Sequence 130, App
c	43	10.8	37.2	45	1	US-08-171-389-342	Sequence 342, App
c	44	10.8	37.2	45	1	US-08-123-936-130	Sequence 130, App
c	45	10.8	37.2	45	1	US-08-123-936-342	Sequence 342, App

ALIGNMENTS

RESULT 1
5258283-10
;Patent No. 5258283
; APPLICANT: FRAZIER, MARVIN E.;MALLAVIA, LOUIS P.;SAMUEL,
;JAMES E.;BACA, OSWALD G.
; TITLE OF INVENTION: DETECTION AND DIFFERENTIATION OF COXIELLA
;BURNETII IN BIOLOGICAL FLUIDS
; NUMBER OF SEQUENCES: 17
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/425,856
; FILING DATE: 23-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 927,779
; FILING DATE: 05-NOV-1986
; APPLICATION NUMBER: 795,207
; FILING DATE: 05-NOV-1985
;SEQ ID NO:10:
; LENGTH: 27
5258283-10

Query Match 42.1%; Score 12.2; DB 6; Length 27;
Best Local Similarity 40.9%; Pred. No. 1.4e+02;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

```
Qy      4 gauncuuunnguaagcccnang 25
          | : |:::   :|||||  | |
Db      4 ggttctttgaataagccaaatg 25
```

RESULT 2
US-08-410-654B-30
; Sequence 30, Application US/08410654B
; Patent No. 5833976
; GENERAL INFORMATION:
; APPLICANT: Rene de Waal Malefyt
; APPLICANT: Di-Hwei Hsu
; APPLICANT: Anne O'Garra
; APPLICANT: Hergen Spits
; TITLE OF INVENTION: Use of Interleukin-10 to Treat
; TITLE OF INVENTION: Septic Shock
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033
; COMPUTER READABLE FORM:

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  Macintosh
;       OPERATING SYSTEM:  7.5.3
;       SOFTWARE:  Microsoft Word 5.1a
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/410,654B
;       FILING DATE:  24-MAR-1995
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/229,854
;       FILING DATE:  19-APR-1994
;       APPLICATION NUMBER:  US 07/926,853
;       FILING DATE:  06-AUG-1992
;       APPLICATION NUMBER:  US 07/742,129
;       FILING DATE:  06-AUG-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Foulke, Cynthia L.
;       REGISTRATION NUMBER:  32,364
;       REFERENCE/DOCKET NUMBER:  DX0221KQ1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  908-298-2987
;       TELEFAX:  908-298-5388
;   INFORMATION FOR SEQ ID NO:  30:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  69 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (oligonucleotide)
US-08-410-654B-30


  Query Match             42.1%;  Score 12.2;  DB 2;  Length 69;
  Best Local Similarity   43.5%;  Pred. No. 1.7e+02;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnangng 27
           |: | ::   :|||| | | | |
Db      11 ATGCCTTTAATAAGCTCCAAGAG 33


RESULT   3
US-08-474-851-30
; Sequence 30, Application US/08474851
; Patent No. 5837232
;  GENERAL INFORMATION:
     APPLICANT:  Rene de Waal Malefyt
     APPLICANT:  Di-Hwei Hsu
     APPLICANT:  Anne O'Garra
;    APPLICANT:  Hergen Spits
;    TITLE OF INVENTION:  Use of An Interleukin-10 Antagonist to Treat
;    TITLE OF INVENTION:  A B Cell Mediated Autoimmune Disorder
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Schering-Plough Corporation
;      STREET:  2000 Galloping Hill Road
;      CITY:  Kenilworth
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07033
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  7.5.3
;      SOFTWARE:  Microsoft Word 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,851
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/410,654
;      FILING DATE:  24-MAR-1995
;      APPLICATION NUMBER:  US 08/229,854
;      FILING DATE:  19-APR-1994
;      APPLICATION NUMBER:  US 07/926,853
;      FILING DATE:  06-AUG-1992
;      APPLICATION NUMBER:  US 07/742,129
;      FILING DATE:  06-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Foulke, Cynthia L.
;      REGISTRATION NUMBER:  32,364
;      REFERENCE/DOCKET NUMBER:  DX0221KQ1GD
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-298-2987
;      TELEFAX:  908-298-5388
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (oligonucleotide)
US-08-474-851-30


  Query Match             42.1%;  Score 12.2;  DB 2;  Length 69;
  Best Local Similarity   43.5%;  Pred. No. 1.7e+02;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnangng 27
           |: | ::   :|||| | | | |
Db      11 ATGCCTTTAATAAGCTCCAAGAG 33


RESULT   4
US-08-481-560-30
; Sequence 30, Application US/08481560
; Patent No. 5837293
;  GENERAL INFORMATION:
;    APPLICANT:  Rene de Waal Malefyt
;    APPLICANT:  Di-Hwei Hsu
;    APPLICANT:  Anne O'Garra
;    APPLICANT:  Hergen Spits
;    TITLE OF INVENTION:  Use of Interleukin-10 to Modulate
;    TITLE OF INVENTION:  Inflammation or T-Cell Mediated
;    TITLE OF INVENTION:  Immune Function
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Schering-Plough Corporation
;      STREET:  2000 Galloping Hill Road
;      CITY:  Kenilworth
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07033
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  7.5.3
;      SOFTWARE:  Microsoft Word 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,560
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/410,654
;      FILING DATE:  24-MAR-1995
;      APPLICATION NUMBER:  US 08/229,854
;      FILING DATE:  19-APR-1994
;      APPLICATION NUMBER:  US 07/926,853
;      FILING DATE:  06-AUG-1992
;      APPLICATION NUMBER:  US 07/742,129
;      FILING DATE:  06-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Foulke, Cynthia L.
;      REGISTRATION NUMBER:  32,364
;      REFERENCE/DOCKET NUMBER:  DX0221KQ1GC
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-298-2987
;      TELEFAX:  908-298-5388
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (oligonucleotide)
US-08-481-560-30


  Query Match             42.1%;  Score 12.2;  DB 2;  Length 69;
  Best Local Similarity   43.5%;  Pred. No. 1.7e+02;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnangng 27
           |: | ::   :|||| | | | |
Db      11 ATGCCTTTAATAAGCTCCAAGAG 33


RESULT   5
US-08-747-536-10
; Sequence 10, Application US/08747536
; Patent No. 5968737
;  GENERAL INFORMATION:
;    APPLICANT:  Ali-Osman, Francis
;    APPLICANT:  Lopez-Berestein, Gabriel
;    APPLICANT:  Buolamwini, John
;    APPLICANT:  Antoun, Gamil
;    APPLICANT:  Lo, Hui-Wen
;    APPLICANT:  Keller, Charles
;    APPLICANT:  Akande, Olanike
;    TITLE OF INVENTION:  GLUTATHIONE S-TRANSFERASE (GST) GENES IN
;    TITLE OF INVENTION:  CANCER
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
       SOFTWARE:  PatentIn Release #1.0, Version #1.30
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/747,536
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTXC:492
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-747-536-10


  Query Match             40.7%;  Score 11.8;  DB 2;  Length 21;
  Best Local Similarity   50.0%;  Pred. No. 2.3e+02;
  Matches    9;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagccc 21
           ||  |:::  |: |||||
Db       2 GAGGCTTTGAGTGAGCCC 19


RESULT   6
US-08-218-369-7/c
; Sequence 7, Application US/08218369
; Patent No. 6312699
;  GENERAL INFORMATION:
;    APPLICANT:  Curiel, David T.
;    APPLICANT:  Engler, Jeffrey A.
;    TITLE OF INVENTION:  Ligands Added to Adenovirus Fiber
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-4530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/218,369
;      FILING DATE:  28-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  IGI101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 815-6508
;      TELEFAX:  (404) 815-6555
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..36
;      OTHER INFORMATION:  /note= "Nucleotide sequence encoding a streptavidin mimic
US-08-218-369-7


  Query Match             40.0%;  Score 11.6;  DB 4;  Length 36;
  Best Local Similarity   41.7%;  Pred. No. 3.4e+02;
  Matches   10;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnangng 27
           ||  |:::  |:  | || | | |
Db      31 GAAGCTTTAGGTGGGGCCCATGAG 8


RESULT   7
US-08-218-369-15
; Sequence 15, Application US/08218369
; Patent No. 6312699
;  GENERAL INFORMATION:
;    APPLICANT:  Curiel, David T.
;    APPLICANT:  Engler, Jeffrey A.
;    TITLE OF INVENTION:  Ligands Added to Adenovirus Fiber
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
```

```
;       STREET:  1100 Peachtree Street, Suite 2800
;       CITY:  Atlanta
;       STATE:  Georgia
;       COUNTRY:  USA
;       ZIP:  30309-4530
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/218,369
;       FILING DATE:  28-MAR-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Pabst, Patrea L.
;       REGISTRATION NUMBER:  31,284
;       REFERENCE/DOCKET NUMBER:  IGI101
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (404) 815-6508
;       TELEFAX:  (404) 815-6555
;   INFORMATION FOR SEQ ID NO:  15:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  36 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  1..36
;       OTHER INFORMATION:  /note= "Nucleotides 5 through 36 are complementary to nucl
US-08-218-369-15


  Query Match             40.0%;  Score 11.6;  DB 4;  Length 36;
  Best Local Similarity   41.7%;  Pred. No. 3.4e+02;
  Matches   10;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||  |:::  |:  | || | | |
Db     10 GAAGCTTTAGGTGGGGCCCATGAG 33


RESULT   8
PCT-US95-03742-7/c
; Sequence 7, Application PC/TUS9503742
;  GENERAL INFORMATION:
;    APPLICANT:  The UAB Research Foundation
;    TITLE OF INVENTION:  Ligands Added to Adenovirus Fiber
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/03742
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  IGI101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..36
;      OTHER INFORMATION:  /note= "Nucleotide sequence
;      OTHER INFORMATION:  encoding a streptavidin mimic that binds biotin."
PCT-US95-03742-7


  Query Match             40.0%;  Score 11.6;  DB 5;  Length 36;
  Best Local Similarity   41.7%;  Pred. No. 3.4e+02;
  Matches   10;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||  |:::  |:  | || | | |
Db     31 GAAGCTTTAGGTGGGGCCCATGAG 8


RESULT   9
PCT-US95-03742-15
; Sequence 15, Application PC/TUS9503742
;  GENERAL INFORMATION:
;    APPLICANT:  The UAB Research Foundation
;    TITLE OF INVENTION:  Ligands Added to Adenovirus Fiber
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/03742
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  IGI101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..36
```

```
;      OTHER INFORMATION:  /note= "Nucleotides 5 through 36
;      OTHER INFORMATION:  are complementary to nucleotides 5 through 36 of
;      OTHER INFORMATION:  Sequence ID No. 7."
PCT-US95-03742-15


  Query Match             40.0%;  Score 11.6;  DB 5;  Length 36;
  Best Local Similarity   41.7%;  Pred. No. 3.4e+02;
  Matches   10;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      4 gauncuuunnguaagcccnangng 27
          ||  |:::  |:  | || | | |
Db     10 GAAGCTTTAGGTGGGGCCCATGAG 33


RESULT  10
US-08-741-881-28/c
; Sequence 28, Application US/08741881
; Patent No. 5789245
;  GENERAL INFORMATION:
     APPLICANT:  Dubensky Jr, Thomas W
     APPLICANT:  Polo, John M.
     APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/741,881
;      FILING DATE:  30-OCT-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
       REFERENCE/DOCKET NUMBER:  930049.423C6 / 1146.007
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-741-881-28


  Query Match             38.6%;  Score 11.2;  DB 1;  Length 25;
  Best Local Similarity   36.4%;  Pred. No. 5.4e+02;
  Matches    8;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      6 uncuuunnguaagcccnangng 27
          : |:::  |: ||||  |   |
Db     24 TCCTTTAGGTTAGCCGTACAAG 3


RESULT  11
US-08-739-158-28/c
; Sequence 28, Application US/08739158
; Patent No. 5814482
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/739,158
;      FILING DATE:  30-OCT-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423D3 / 1146.012
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-739-158-28


  Query Match             38.6%;  Score 11.2;  DB 1;  Length 25;
  Best Local Similarity   36.4%;  Pred. No. 5.4e+02;
  Matches    8;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      6 uncuuunnguaagcccnangng 27
          : |:::  |: ||||  |   |
Db     24 TCCTTTAGGTTAGCCGTACAAG 3


RESULT  12
US-08-739-167-28/c
; Sequence 28, Application US/08739167
; Patent No. 5843723
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS  AND ALPHAVIRU
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/739,167
;       FILING DATE:  30-OCT-1996
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  McMasters, David D.
;       REGISTRATION NUMBER:  33,963
;       REFERENCE/DOCKET NUMBER:  930049.423C7 / 1146.008
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  28:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  25 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
        TOPOLOGY:  linear
    08-739-167-28


  Query Match             38.6%;  Score 11.2;  DB 2;  Length 25;
  Best Local Similarity   36.4%;  Pred. No. 5.4e+02;
  Matches    8;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       6 uncuuunnguaagcccnangng 27
           : |:::  |: ||||  |   |
Db      24 TCCTTTAGGTTAGCCGTACAAG 3


RESULT  13
US-08-404-796-28/c
; Sequence 28, Application US/08404796
; Patent No. 6015686
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/404,796
;      FILING DATE:  15-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C5 / 1146.006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-404-796-28


  Query Match             38.6%;  Score 11.2;  DB 3;  Length 25;
  Best Local Similarity   36.4%;  Pred. No. 5.4e+02;
  Matches    8;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       6 uncuuunnguaagcccnangng 27
           : |:::  |: ||||  |   |
Db      24 TCCTTTAGGTTAGCCGTACAAG 3


RESULT  14
US-08-931-869-28/c
; Sequence 28, Application US/08931869
; Patent No. 6015694
;  GENERAL INFORMATION:
;    APPLICANT:  Dubensky Jr, Thomas W
;    APPLICANT:  Polo, John M.
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Driver, David A.
;    APPLICANT:  Belli, Barbara A.
;    TITLE OF INVENTION:  EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;    NUMBER OF SEQUENCES:  128
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,869
;      FILING DATE:  16-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/404,796
;      FILING DATE:  15-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  930049.423C5 / 1146.006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-931-869-28


  Query Match             38.6%;  Score 11.2;  DB 3;  Length 25;
  Best Local Similarity   36.4%;  Pred. No. 5.4e+02;
  Matches    8;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       6 uncuuunnguaagcccnangng 27
           : |:::  |: ||||  |   |
Db      24 TCCTTTAGGTTAGCCGTACAAG 3


RESULT  15
US-09-350-399-28/c
```

```
; Sequence 28, Application US/09350399
; Patent No. 6342372
;   GENERAL INFORMATION:
;        APPLICANT: Dubensky Jr, Thomas W
;                   Polo, John M.
;                   Jolly, Douglas J.
;                   Driver, David A.
;        TITLE OF INVENTION: EUKARYOTIC LAYERED VECTOR INITIATION SYSTEMS
;        NUMBER OF SEQUENCES: 128
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: SEED and BERRY LLP
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: US
;             ZIP: 98104-7092
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/350,399
;             FILING DATE: 08-Jul-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McMasters, David D.
;             REGISTRATION NUMBER: 33,963
;             REFERENCE/DOCKET NUMBER: 930049.423D1 / 1146.010
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 28:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 25 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-350-399-28


  Query Match           38.6%;  Score 11.2;  DB 4;  Length 25;
  Best Local Similarity 36.4%;  Pred. No. 5.4e+02;
  Matches    8;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       6 uncuuunnguaagcccnangng 27
           : |:::  |: ||||  |   |
Db      24 TCCTTTAGGTTAGCCGTACAAG 3


Search completed: April 29, 2002, 22:39:48
Job time: 10159 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 29, 2002, 18:34:27 ; Search time 1874.63 Seconds
(without alignments)
208.794 Million cell updates/sec

Title: US-09-310-844C-23
Perfect score: 29
Sequence: 1 nnngauncuuunnguaagcccnangngnn 29

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 297742

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	13	44.8	46	12	AZ833686	AZ833686 2M0115L20
c	2	12.8	44.1	76	12	AQ025263	AQ025263 EP(3)3081
	3	12.8	44.1	86	12	CNS02IOD	AL199174 Tetraodon
c	4	12.4	42.8	52	9	AA700959	AA700959 zf87d10.s
c	5	12.4	42.8	70	9	AA468615	AA468615 ne08b04.s
	6	12.2	42.1	48	12	AZ503560	AZ503560 1M0343E21
c	7	12.2	42.1	75	12	AZ453746	AZ453746 1M0255A23
c	8	12.2	42.1	86	10	C01512	C01512 HUMGS000850
	9	12.2	42.1	92	10	BI305219	BI305219 NRS2-7_8_
	10	12.2	42.1	96	10	H81976	H81976 yv84e12.r1
c	11	12.2	42.1	98	12	AZ767813	AZ767813 1M0567G24
c	12	11.8	40.7	40	9	AA975071	AA975071 on03d07.s
c	13	11.8	40.7	49	10	BE970036	BE970036 601680150
c	14	11.8	40.7	65	9	AA733449	AA733449 vt73h08.r
	15	11.8	40.7	70	9	AI767928	AI767928 wi99c01.x
c	16	11.8	40.7	70	12	BH216023	BH216023 1006039G0
c	17	11.8	40.7	75	12	CNS01561	AL105043 Drosophil
c	18	11.8	40.7	83	10	BE845147	BE845147 AD07E12T7
c	19	11.8	40.7	86	10	BE845145	BE845145 AD07E06T7
	20	11.8	40.7	87	12	AZ807716	AZ807716 2M0070D13
	21	11.8	40.7	93	9	AA488835	AA488835 aa54h10.r
c	22	11.8	40.7	95	12	AZ598556	AZ598556 1M0413K23
c	23	11.8	40.7	98	9	AW311302	AW311302 sg35b11.y
	24	11.6	40.0	49	10	U44334	U44334 ENU44334 As
c	25	11.6	40.0	58	9	AI584456	AI584456 fb93h12.x
	26	11.6	40.0	70	9	AI814489	AI814489 wj73g11.x
c	27	11.6	40.0	79	12	AZ336769	AZ336769 1M0067L12
	28	11.6	40.0	81	12	BH627770	BH627770 1007076E0
c	29	11.6	40.0	85	9	AA617776	AA617776 np99e08.s
c	30	11.6	40.0	85	10	W20254	W20254 zb42a10.r1
	31	11.6	40.0	93	12	AZ832088	AZ832088 2M0112D15
c	32	11.6	40.0	94	9	AA558052	AA558052 nl18f06.s
c	33	11.6	40.0	97	9	AI181481	AI181481 uc62g11.r
c	34	11.6	40.0	100	9	AI148625	AI148625 qc62g05.x
c	35	11.6	40.0	100	10	H13996	H13996 EST00022 Ch
c	36	11.4	39.3	32	12	HSMC42B09	X88068 H.sapiens D
c	37	11.4	39.3	40	12	TA253H01Q	AL483109 T. brucei
	38	11.4	39.3	46	9	AI887082	AI887082 wl96e09.x
c	39	11.4	39.3	66	9	AA247859	AA247859 j3371.seq
	40	11.4	39.3	67	12	TA113E04Q	AL460350 T. brucei
c	41	11.4	39.3	73	9	AI900474	AI900474 sc11b10.y
c	42	11.4	39.3	75	12	BH609927	BH609927 HIV22C11
	43	11.4	39.3	76	9	AA988261	AA988261 os16a07.s
	44	11.4	39.3	81	9	AA738756	AA738756 vv67b08.r
c	45	11.4	39.3	85	12	AZ830379	AZ830379 2M0109A09

ALIGNMENTS

RESULT 1
AZ833686
LOCUS AZ833686 46 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0115L20R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0115L20 R, DNA sequence.
ACCESSION AZ833686
VERSION AZ833686.1 GI:13003594
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 46)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0115 row: L column: 20
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 46.
FEATURES Location/Qualifiers
source 1. .46
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0115L20"
/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 14 a 6 c 13 g 13 t
ORIGIN

Query Match 44.8%; Score 13; DB 12; Length 46;
Best Local Similarity 50.0%; Pred. No. 1.4e+04;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 4 gauncuuunnguaagc 19
||: |::: |:|||
Db 8 GATACTTTAAGTAAGC 23

RESULT 2
AQ025263/c
LOCUS AQ025263 76 bp DNA linear GSS 23-AUG-2000
DEFINITION EP(3)3081 Drosophila melanogaster EP line Drosophila melanogaster genomic Sequence recovered from 5' end of P element, DNA sequence.
ACCESSION AQ025263
VERSION AQ025263.1 GI:3265615
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 76)
AUTHORS Liao,G.-C., Rehm,E.J. and Rubin,G.M.
TITLE Insertion site preferences of the P transposable element in Drosophila melanogaster
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
MEDLINE 20202638
COMMENT Contact: Gerald Rubin
Berkeley Drosophila Genome Project
University of California, Berkeley
LSA Building, Berkeley, CA 94720-3200, USA
Fax: 5106439947
Email: gerry@fruitfly.berkeley.edu
Sequence recovery method was inverse PCR.

Sequence orientation is forward strand relative to 5' end of P element

The P element insertion position is base 69 in the 76 bases. This insertion position refers to the first base of the 8 base target recognition sequence.
Class: transposon-tagged.
FEATURES Location/Qualifiers
source 1. .76
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="Drosophila melanogaster EP line"
/note="Inverse PCR was performed on Drosophila melanogaster strains each of which contains a single EP transposable element insertion. (The generation of these insertion strains is described in Rorth P, Szabo K, Bailey A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function genetics in Drosophila. Development 6:1049-1057.) The resultant fragment for each strain was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."
BASE COUNT 31 a 11 c 13 g 21 t
ORIGIN

Query Match 44.1%; Score 12.8; DB 12; Length 76;
Best Local Similarity 42.9%; Pred. No. 1.8e+04;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 5 auncuuunnguaagcccnang 25
|: |::: :|| ||| | |
Db 62 ATACTTTATTTAATCCCAAAG 42

RESULT 3
CNS02IOD
LOCUS CNS02IOD 86 bp DNA linear GSS 13-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 141P20 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL199174
VERSION AL199174.1 GI:7837325
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 86)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Charaterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 86)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 86)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.
FEATURES Location/Qualifiers
source 1. .86
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="141P20"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG141DH10LP1-end : T7"
BASE COUNT 12 a 12 c 26 g 34 t 2 others
ORIGIN

Query Match 44.1%; Score 12.8; DB 12; Length 86;

```
  Best Local Similarity   40.9%;  Pred. No. 1.9e+04;
  Matches    9;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       6 uncuuunnguaagcccnangng 27
           : |:::  |:| | || | | |
Db      62 TTCTTTTTGTAGGGCCTAGGGG 83


RESULT   4
AA700959/c
LOCUS       AA700959                 52 bp    mRNA    linear   EST 19-DEC-1997
DEFINITION  zf87d10.s1 Soares_pineal_gland_N3HPG Homo sapiens cDNA clone
            IMAGE:383923 3' similar to TR:P79324 P79324 RIBOSOMAL PROTEIN L15
            ;, mRNA sequence.
ACCESSION   AA700959
VERSION     AA700959.1  GI:2704124
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Homo sapiens"
                     /db_xref="GDB:1292180"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:383923"
                     /clone_lib="Soares_pineal_gland_N3HPG"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: pineal gland; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCGTTTTTTTTTTTTTTTTTT 3']
                     , double-stranded cDNA was size selected, ligated to Eco
                     RI adapters (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT       17 a     12 c     12 g     11 t
ORIGIN


  Query Match             42.8%;  Score 12.4;  DB 9;  Length 52;
  Best Local Similarity   52.9%;  Pred. No. 2.9e+04;
  Matches    9;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcc 20
           ||: | ::  |:|||||
Db      31 GATCCATTCCGTAAGCC 15


RESULT   5
AA468615/c
LOCUS       AA468615                 70 bp    mRNA    linear   EST 15-AUG-1997
DEFINITION  ne08b04.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:880591 3'
            similar to TR:G178246 G178246 MATURE ALPHA-GALACTOSIDASE A,
            COMPLETE CDS PRECURSOR ;, mRNA sequence.
ACCESSION   AA468615
VERSION     AA468615.1  GI:2195149
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.
            , Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1285   Std Error: 0.00
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:880591"
                     /clone_lib="NCI_CGAP_Co3"
                     /sex="pooled"
                     /tissue_type="colon"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from 12 pooled bulk tumor samples and primed
                     with a Not I - oligo(dT) primer. Double-stranded cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector. Library went through one round of
                     normalization. "
BASE COUNT       17 a     22 c     14 g     17 t
ORIGIN


  Query Match             42.8%;  Score 12.4;  DB 9;  Length 70;
  Best Local Similarity   52.9%;  Pred. No. 2.9e+04;
  Matches    9;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcc 20
           ||: |:::  | |||||
Db      64 GATCCTTTGGGCAAGCC 48


RESULT   6
AZ503560
LOCUS       AZ503560                 48 bp    DNA     linear   GSS 05-OCT-2000
DEFINITION  1M0343E21F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0343E21 F, DNA sequence.
ACCESSION   AZ503560
VERSION     AZ503560.1  GI:10684876
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
```

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0343 row: E column: 21
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 48.
TURES Location/Qualifiers
source 1. .48
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0343E21"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
SE COUNT 14 a 7 c 17 g 10 t
GIN

Query Match 42.1%; Score 12.2; DB 12; Length 48;
Best Local Similarity 43.5%; Pred. No. 3.6e+04;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 5 auncuuunnguaagcccnangng 27
|: |: : |:||| | | | |
Db 13 ATGCTCTGGGTAAGGCACAAGAG 35

RESULT 7
AZ453746/c
LOCUS AZ453746 75 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0255A23F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0255A23 F, DNA sequence.
ACCESSION AZ453746
VERSION AZ453746.1 GI:10611850
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 75)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0255 row: A column: 23
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 75.
FEATURES Location/Qualifiers
source 1. .75
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0255A23"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 22 a 16 c 23 g 14 t
ORIGIN

Query Match 42.1%; Score 12.2; DB 12; Length 75;
Best Local Similarity 43.5%; Pred. No. 3.8e+04;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 5 auncuuunnguaagcccnangng 27
|: : : |:||| || | | |
Db 24 ATAATGTTTGTAAGTCCAATGGG 2

RESULT 8
C01512/c
LOCUS C01512 86 bp mRNA linear EST 23-JUL-1996
DEFINITION HUMGS0008507 Human adult (K.Okubo) Homo sapiens cDNA, mRNA sequence.
ACCESSION C01512
VERSION C01512.1 GI:1433742
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 86)

```
  AUTHORS   Okubo,K.
  TITLE     BodyMap; human gene expression database
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Okubo,K.
            Institute for Molecular and Cellular Biol
            Osaka University
            1-3,Yamada-oka, Suita, Osaka Pref. 565, Japan
            Tel: 06-877-5111(ex.3315)
            Email: kousaku@imcb.osaka-u.ac.jp
            Human Gene Signature, 3'-directed cDNA sequence. We are not
            submitting the same cDNA sequence redundantly to DDBJ since 1993.
            For the abundance information of clones with this sequence in this
            library and as well as in other 3'-directed libraries, see '
            http://www.imcb.osaka-u.ac.jp/bodymap'. The sequences of the clones
            represented by this GS sequences is also found there.
FEATURES             Location/Qualifiers
     source          1. .86
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Human adult (K.Okubo)"
                     /dev_stage="adult"
   E COUNT       25 a     21 c     15 g     24 t      1 others
   GIN


  Query Match             42.1%;  Score 12.2;  DB 10;  Length 86;
  Best Local Similarity   47.8%;  Pred. No. 3.8e+04;
  Matches   11;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnangng 27
           |:| :::  | |||||  | | |
Db      64 ATNATTTCAGAAAGCCAAAGGGG 42
;
RESULT   9
BI305219
LOCUS       BI305219                  92 bp    mRNA    linear   EST 20-JUL-2001
DEFINITION  NRS2-7_8_O20 Drought stress (root) Oryza sativa cDNA clone
            NRS2-7_8_O20 3', mRNA sequence.
ACCESSION   BI305219
VERSION     BI305219.1  GI:14980541
KEYWORDS    EST.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   Reddy,A.R., Ramakrishna,W., Chandrasekhar,A., Nagabhushan,I.,
            Ravindrababu,P. and Bennetzen,J.L.
  'ITLE     Novel EST enrichment with normalized cDNA libraries from drought
            stressed rice (Oryza sativa L.cv Nagina 22)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Reddy AR
            Department of Plant Sciences, School of Life Sciences
            University of Hyderabad
            P.O. Central University, Hyderabad-500 046, A.P, India
            Tel: 0091-40-3010265
            Fax: 0091-40-3010145
            Email: arjulsl@uohyd.ernet.in
            Insert Length: 92   Std Error: 0.00
            Plate: 8  row: 20  column: O
            Seq primer: GTAAAACGACGGCCAGTG.
FEATURES             Location/Qualifiers
     source          1. .92
                     /organism="Oryza sativa"
                     /cultivar="Nagina 22 (indica sub sp)"
                     /db_xref="taxon:4530"
                     /clone="NRS2-7_8_O20"
                     /clone_lib="Drought stress (root)"
                     /tissue_type="Entire root tissue"
                     /dev_stage="35 day-old seedlings"
                     /note="Organ: Leaf; Vector: T7T3Pac; ESTs from normalized
                root cDNA Library from drought stressed seedlings "
BASE COUNT       18 a     19 c     16 g     39 t
ORIGIN


  Query Match             42.1%;  Score 12.2;  DB 10;  Length 92;
  Best Local Similarity   43.5%;  Pred. No. 3.8e+04;
  Matches   10;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnangng 27
           |  |:::  |:| || | | | |
Db       8 AAGCTTTATGTATGCTCAACGCG 30


RESULT  10
H81976
LOCUS       H81976                    96 bp    mRNA    linear   EST 09-NOV-1995
DEFINITION  yv84e12.r1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
            IMAGE:249454 5' similar to gb:M64241 QM PROTEIN (HUMAN);, mRNA
            sequence.
ACCESSION   H81976
VERSION     H81976.1  GI:1060065
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Seq primer: M13RP1
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Homo sapiens"
                     /db_xref="GDB:3867160"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:249454"
                     /clone_lib="Soares melanocyte 2NbHM"
                     /sex="Male"
                     /tissue_type="melanocyte"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. RNA from normal foreskin melanocytes
                     (FS374) was kindly provided by Dr. Anthony P. Albino."
BASE COUNT       17 a     23 c     26 g     17 t     13 others
ORIGIN


  Query Match             42.1%;  Score 12.2;  DB 10;  Length 96;
```

```
  Best Local Similarity   50.0%;  Pred. No. 3.8e+04;
  Matches   11;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagcccnang 25
           | : | ::  | ||||| | |
Db      72 GGTGCCTTTGGAAAGCCCCAGG 93


RESULT  11
AZ767813/c
LOCUS       AZ767813                98 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  1M0567G24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0567G24 F, DNA sequence.
ACCESSION   AZ767813
VERSION     AZ767813.1  GI:12886296
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0567  row: G  column: 24
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 98.
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0567G24"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       17 a     32 c     22 g     27 t
ORIGIN


  Query Match             42.1%;  Score 12.2;  DB 12;  Length 98;
  Best Local Similarity   47.8%;  Pred. No. 3.8e+04;
  Matches   11;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccnangng 27
           |: | :   |:||| || | | |
Db      73 ATCCCTACAGTAAGACCTAGGGG 51


RESULT  12
AA975071/c
LOCUS       AA975071                40 bp    mRNA    linear   EST 26-AUG-1998
DEFINITION  on03d07.s1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1555597 3'
            similar to TR:P70566 P70566 N-TROPOMODULIN. ;, mRNA sequence.
ACCESSION   AA975071
VERSION     AA975071.1  GI:3150863
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 2096   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1555597"
                     /clone_lib="NCI_CGAP_Kid3"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer,
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia),  digested with Not I and cloned into the Not
                     I and Eco RI sites  of the modified pT7T3 vector.  mRNA
                     source: 2 pooled kidneys. Library went through one round
                     of normalization. Library constructed by Bento Soares and
                     M. Fatima Bonaldo.  "
BASE COUNT       10 a      7 c     11 g     12 t
ORIGIN


  Query Match             40.7%;  Score 11.8;  DB 9;  Length 40;
  Best Local Similarity   40.0%;  Pred. No. 5.8e+04;
  Matches    8;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6 uncuuunnguaagcccnang 25
           : |::  |:||| ||   |
Db      35 TCCTTTCCGTAAGACCTTGG 16


RESULT  13
BE970036/c
LOCUS       BE970036                49 bp    mRNA    linear   EST 04-OCT-2000
```

```
DEFINITION  601680150F1 NIH_MGC_78 Homo sapiens cDNA clone IMAGE:3950172 5',
            mRNA sequence.
ACCESSION   BE970036
VERSION     BE970036.1  GI:10582969
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM816  row: d  column: 13
            High quality sequence stop: 49.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3950172"
                     /clone_lib="NIH_MGC_78"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: pancreas; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                     ); 5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.2 kb (range 0.5-4.0 kb).  14/15 colonies
                     contained inserts  by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT       20 a      7 c     12 g     10 t
ORIGIN
```

```
  Query Match            40.7%;  Score 11.8;  DB 10;  Length 49;
  Best Local Similarity  45.0%;  Pred. No. 5.9e+04;
  Matches    9;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       6 uncuuunnguaagcccnang 25
           :  :::  | ||||||  | |
Db      48 TTTTTTATGCAAGCCCCAGG 29
```

```
RESULT  14
AA733449/c
LOCUS       AA733449                 65 bp    mRNA    linear   EST 07-JAN-1998
DEFINITION  vt73h08.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone IMAGE:1176831 5' similar to gb:X06617 40S RIBOSOMAL PROTEIN
            S11 (HUMAN);, mRNA sequence.
ACCESSION   AA733449
VERSION     AA733449.1  GI:2755116
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:634679
            Trace considered overall poor quality
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1176831"
                     /clone_lib="Barstead mouse irradiated colon MPLRB7"
                     /dev_stage="8 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                     from 8 week old mouse.  Colon was harvested 72 hours after
                     irradiation with 1400 Gys.  1st strand cDNA was primed
                     with  a Not I - oligo(dT) primer
                     [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors [AATTCGGATCCTTG], digested with Not I and cloned
                     into the Not I and Eco RI sites of  the modified pT7T3
                     vector. Library constructed by Bob Barstead.  "
BASE COUNT       23 a     15 c     16 g     11 t
ORIGIN
```

```
  Query Match            40.7%;  Score 11.8;  DB 9;  Length 65;
  Best Local Similarity  44.4%;  Pred. No. 6e+04;
  Matches    8;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       4 gauncuuunnguaagccc 21
           | : |:::  |:|||| |
Db      44 GCTGCTTTTGGTAAGCAC 27
```

```
RESULT  15
AI767928
LOCUS       AI767928                 70 bp    mRNA    linear   EST 21-DEC-1999
DEFINITION  wi99c01.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2401440 3'
            similar to SW:RT14_HUMAN P78537 RT14 PROTEIN ;, mRNA sequence.
ACCESSION   AI767928
VERSION     AI767928.1  GI:5234426
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
```

```
           www-bio.llnl.gov/bbrp/image/image.html

           Trace considered overall poor quality
           Insert Length: 574    Std Error: 0.00
           Seq primer: -40UP from Gibco
           High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2401440"
                     /clone_lib="NCI_CGAP_Kid12"
                     /tissue_type="2 pooled tumors (clear cell type)"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1323912-1325831, 1471368-1472903 and
                     1492104-1493255). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
BASE COUNT       20 a      24 c       6 g      20 t
ORIGIN


  Query Match             40.7%;  Score 11.8;  DB 9;  Length 70;
  Best Local Similarity   47.4%;  Pred. No. 6e+04;
  Matches    9;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 auncuuunnguaagcccna 23
           |: |:::  | |||||  |
Db      46 ATTCTTTAAGCAAGCCAGA 64


Search completed: April 29, 2002, 22:38:40
Job time: 14653 sec
```